MEME16203Linear Models Marking Guide

Assignment 4

UNIVERSITI TUNKU ABDUL RAHMAN

Faculty:	FES	Unit Code:	MEME16203
Course:	MAC	Unit Title:	Linear Models
Year:	1,2	Lecturer:	Dr Yong Chin Khian
Session:	May 2022	Due by:	09/07/2022

Q1. An researcher recorded moisture content for three types of cheese made by two different methods. Three pieces of cheese were measure for each type and each method. The data are shown below.

Treatment	Moisture Content Measurements		
Type A made with Method 1	$y_{11} = 38.02$	$y_{12} = 39.79$	$y_{13} = 37.79$
Type B made with Method 1	$y_{21} = 36.74$	$y_{22} = 33.41$	$y_{23} = 38.41$
Type C made with Method 1	$y_{31} = 38.02$	$y_{32} = 35.00$	$y_{33} = 34.00$
Type A made with Method 2	$y_{41} = 39.96$	$y_{42} = 39.06$	$y_{43} = 38.01$
Type B made with Method 2	$y_{51} = 34.58$	$y_{52} = 36.52$	$y_{53} = 35.52$
Type C made with Method 2	$y_{61} = 34.60$	$y_{62} = 36.05$	$y_{63} = 38.0$

Consider the model $y_{ij} = \mu + \alpha_i + \epsilon_{ij}$, where $\epsilon_{ij} \sim NID(0, \sigma^2)$, $i = 1, 2, 3, 4, 5, 6$, and $j = 1, 2, 3$. This model can be expressed in matrix form as $\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\epsilon}$.

Express the each of the following hypotheses in the form $H_0 : \mathbf{C}\boldsymbol{\beta} = \mathbf{0}$. If the hypothesis is testable, compute the value of the corresponding F-statistic and report the degrees of freedom.

(a) After averaging across the two methods of making cheese, the average moisture content is the same for all three types of cheese. [Note: SSE = 33.72387 and $SSH_0 = 32.91023$] (10 marks)

Ans.

The hypothesis to be tested is

$$H_0 : (\alpha_1 + \alpha_4)/2 = (\alpha_2 + \alpha_5)/2 = (\alpha_3 + \alpha_6)/2$$

or

$$H_0 : \mathbf{C}\boldsymbol{\beta} = \begin{bmatrix} 0 & 1 & -1 & 0 & 1 & -1 & 0 \\ 0 & 1 & 0 & -1 & 1 & 0 & -1 \end{bmatrix} \boldsymbol{\beta} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

$F = \frac{MSH_0}{MSE} = \frac{32.91023/2}{33.72387/12} = 5.8552$ with $df = (2, 12)$

(b) For each type of cheese, the average moisture content is not affected by the method for making cheese. (This hypothesis allows the average moisture content to vary across types of cheese).[Note: $SSH_0 = 1.4109$] (10 marks)

Ans. The hypothesis to be tested is

$$H_0 : \alpha_1 = \alpha_4; \alpha_2 = \alpha_5; \alpha_3 = \alpha_6$$

or

$$H_0 : \mathbf{C}\boldsymbol{\beta} = \begin{bmatrix} 0 & 1 & 0 & 0 & -1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & -1 \end{bmatrix} \boldsymbol{\beta} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

$F = \frac{MSH_0}{MSE} = \frac{1.4109/3}{33.72387/12} = 0.16735$ with $df = (3, 12)$

<u>MEME16203Linear Models Marking Guide</u>

Q2. Let $\mathbf{Y} \sim N(\mathbf{W}\boldsymbol{\gamma}, \sigma^2 I)$, where

- $\mathbf{W} = \left[\mathbf{W_1}\ \mathbf{W_2}\ \mathbf{W_3}\ \mathbf{W_4}\right]$,
- $\mathbf{W_1} = \mathbf{1_{20}}$,
- $\mathbf{W_2} = \begin{bmatrix} 1 \\ -1 \end{bmatrix} \otimes \mathbf{1_{10}}$,
- $\mathbf{W_3} = \mathbf{1_2} \otimes \begin{bmatrix} 1 \\ -1 \end{bmatrix} \otimes \mathbf{1_5}$,
- $\mathbf{W_4} = \mathbf{1_4} \otimes \begin{bmatrix} -4 \\ -2 \\ 0 \\ 4 \\ 2 \end{bmatrix}$, and
- $\boldsymbol{\gamma} = \begin{bmatrix} \gamma_1 \\ \gamma_2 \\ \gamma_3 \\ \gamma_4 \end{bmatrix}$

(a) Use Cochran's theorem to find the distributions of

- $\frac{1}{\sigma^2} SSE = \mathbf{e^T e} = \mathbf{Y^T(I - P_W)Y}$, where $\mathbf{P_W} = \mathbf{W(W^TW)^{-1}W^T}$
- $\frac{1}{\sigma^2} R(\gamma_1) = \mathbf{Y^T P_{W_1} Y}$ where $\mathbf{W_1} = \mathbf{1}$ is the first column of $\mathbf{W}$ and $\mathbf{P_{W_1}} = \mathbf{W_1(W_1^T W)^{-1} W_1^T}$.
- $\frac{1}{\sigma^2} R(\gamma_2|\gamma_1) = \mathbf{Y^T}(\mathbf{P_{W_2}} - \mathbf{P_{W_1}})Y$ where $\mathbf{W_2}$ contains the first two columns of $\mathbf{W}$ and $\mathbf{P_{W_2}} = \mathbf{W_2(W_2^T W_2)^{-1} W_2^T}$.
- $\frac{1}{\sigma^2} R(\gamma_3|\gamma_1\gamma_2) = \mathbf{Y^T}(\mathbf{P_{W_3}} - \mathbf{P_{W_2}})\mathbf{Y}$. where $\mathbf{W_3}$ contains the first three columns of $\mathbf{W}$ and $\mathbf{P_{W_3}} = \mathbf{W_3(W_3^T W_3)^{-1} W_3^T}$.
- $\frac{1}{\sigma^2} R(\gamma_4|\gamma_1\gamma_2\gamma_3) = \mathbf{Y^T}(\mathbf{P_W} - \mathbf{P_{W_3}})\mathbf{Y}$.

(10 marks)

Ans.

Check the conditions of Cochran's Theorem.

1) $(\mathbf{I} - \mathbf{P_W}) + \mathbf{P_{W_1}} + (\mathbf{P_{W_2}} - \mathbf{P_{W1}}) + (\mathbf{P_{W_3}} - \mathbf{P_{W_2}}) + (\mathbf{P_W} - \mathbf{P_{W_3}}) = \mathbf{I}$

2) Rank$(\mathbf{I} - \mathbf{P_W})$+Rank$\mathbf{P_{W_1}}$ + Rank$(\mathbf{P_{W_2}} - \mathbf{P_{W1}})$ +Rank$(\mathbf{P_{W_3}} - \mathbf{P_{W_2}}$ + Rank$(\mathbf{P_W} - \mathbf{P_{W_3}}) = (20-4)+1+(2-1)+(3-2)+(4-3) = 20 = n$

3) $(\mathbf{I} - \mathbf{P_W})$, $\mathbf{P_{W_1}}$, $(\mathbf{P_{W_2}} - \mathbf{P_{W1}})$, $(\mathbf{P_{W_3}} - \mathbf{P_{W_2}})$, $(\mathbf{P_W} - \mathbf{P_{W_3}})$ are all symmetric.

Therefore, by Cochran's theorem, the sums of squares are independently distributed as chi-square random, variables with d.f. 16, 1, 1, 1, 1 respectively.

(b) Report a formula for the non-centrality parameter of the non-central F distribution of

$$F = \frac{R(\gamma_3|\gamma_1, \gamma_2)}{SSE/7}$$

Use it to the null and alternative hypotheses associated with this test statistic. You are given that: $\mathbf{W^T}(\mathbf{P_{W_3}} - \mathbf{P_{W_2}})\mathbf{W}] = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 20 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$. (10 marks)

Ans.

$$\lambda = \frac{1}{\sigma^2}\boldsymbol{\beta}^{\mathbf{T}}\mathbf{W^T}(P_{W_3} - P_{W_2})\mathbf{W}\boldsymbol{\beta} = \frac{1}{\sigma^2}\begin{bmatrix}\gamma_1 & \gamma_2 & \gamma_3 & \gamma_4\end{bmatrix}\begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 20 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}\begin{bmatrix}\gamma_1 \\ \gamma_2 \\ \gamma_3 \\ \gamma_4\end{bmatrix} = \frac{20}{\sigma^2}\gamma_3^2$$

Therefore, $\lambda = 0$ iff $\gamma_3 = 0$ and the hypothesis is $H_0 : \gamma_3 = 0$ against the alternative $H_1 : \gamma_3 \neq 0$.

Q3. A chemical production process consists of a first reaction with an alcohol and a second reaction with a base. A 3×2 factorial experiment with three alcohols and two bases was conducted. The data had unequal replications among the six treatment combinations of the two factors, Base and Alcohol. The collected data are percent yield. The data are given below.

	Alcohol		
Base	1	2	3
1	90.3 91.3	89.7 88.3 90.0	89.9 87.0 89.4 90.5
2	88.4 89.1 91.5	95.7	94.8 92.3 91.8

Consider the model $y_{ijk} = \mu_{ij} + \epsilon_{ijk}$, where $\epsilon_{ijk} \sim NID(0, \sigma^2)$, $i = 1, 2$, and $j = 1, 2, 3$ and $k = 1, \ldots, n_{ij}$. This model can be expressed in matrix form as $\mathbf{Y} = \mathbf{D}\boldsymbol{\beta} + \boldsymbol{\epsilon}$. Examine type III sums of squares for these data.

(a) Specify the $\mathbf{C}$ matrix needed to write the null hypothesis associated with the F-test for Base effects in the form $H_0 : \mathbf{C_1}\boldsymbol{\beta} = \mathbf{0}$. (10 marks)

Ans.

$$\mathbf{C_1} = [\mathbf{I}_{a-1}|-\mathbf{1}_{a-1}]\otimes\mathbf{1}_{\mathbf{b}}^{\mathbf{T}} = [\mathbf{I}_1|-\mathbf{1}_1]\otimes\mathbf{1}_{\mathbf{3}}^{\mathbf{T}} = \begin{bmatrix}1 & -1\end{bmatrix}\otimes\begin{bmatrix}1 & 1 & 1\end{bmatrix} = \begin{bmatrix}1 & 1 & 1 & -1 & -1 & -1\end{bmatrix}$$

(b) Present a formula for $SS_{H_{0,1}}$, corresponding to the null hypothesis in part (a), and state it's distribution when the null hypothesis is true.(10 marks)

Ans.

$$\begin{aligned} SS_{H_{0,1}} &= (\mathbf{C_1 b} - \mathbf{0})^{\mathbf{T}}[\mathbf{C_1}(\mathbf{D^T D})^{-1}\mathbf{C_1^T}]^{-1}(\mathbf{C_1 b} - \mathbf{0}) \\ &= \mathbf{y^T D}(\mathbf{D^T D})^{-1}\mathbf{C_1^T}[\mathbf{C_1}(\mathbf{D^T D})^{-1}\mathbf{C_1^T}]^{-1}\mathbf{C_1}(\mathbf{D^T D})^{-1}\mathbf{D^T y} \end{aligned}$$

Under H_0, $SSH_{0,1} \sim \frac{1}{\sigma^2}\chi_1^2$

(c) Compute $SS_{H_{0,1}}$. (10 marks)

Ans.

$$\mathbf{b} = \begin{bmatrix}\bar{Y}_{11.} & \bar{Y}_{12.} & \bar{Y}_{13.} & \bar{Y}_{21.} & \bar{Y}_{22.} & \bar{Y}_{23.}\end{bmatrix}^T = \begin{bmatrix}90.8 & 89.333 & 89.2 & 89.667 & 95.7 & 92.967\end{bmatrix}^T$$

$$\mathbf{C_1 b} = \begin{bmatrix}1 & 1 & 1 & -1 & -1 & -1\end{bmatrix}\begin{bmatrix}90.8\\89.333\\89.2\\89.667\\95.7\\92.967\end{bmatrix} = \begin{bmatrix}-9.001\end{bmatrix}$$

$$\mathbf{D^T D^{-1}} = diag(\tfrac{1}{2}, \tfrac{1}{3}, \tfrac{1}{4}, \tfrac{1}{3}, 1, \tfrac{1}{3})$$

$$C_1(D^TD)^{-1}C^T] = \begin{bmatrix}1 & 1 & 1 & -1 & -1 & -1\end{bmatrix}\begin{bmatrix}\frac{1}{2} & 0 & 0 & 0 & 0 & 0\\ 0 & \frac{1}{3} & 0 & 0 & 0 & 0\\ 0 & 0 & \frac{1}{4} & 0 & 0 & 0\\ 0 & 0 & 0 & \frac{1}{3} & 0 & 0\\ 0 & 0 & 0 & 0 & 1 & 0\\ 0 & 0 & 0 & 0 & 0 & \frac{1}{3}\end{bmatrix}\begin{bmatrix}1\\1\\1\\-1\\-1\end{bmatrix}$$
$$= \begin{bmatrix}2.75\end{bmatrix}$$

$$SSH_{01} = (\mathbf{C_1b-0})^\mathbf{T}[\mathbf{C_1(D^TD)^{-1}C_1^T}]^{-1}(\mathbf{C_1b-0}) = \begin{bmatrix}-9.001\end{bmatrix}\begin{bmatrix}0.3636\end{bmatrix}\begin{bmatrix}-9.001\end{bmatrix} = 29.4611$$

(d) Specify the $\mathbf{C_2}$ matrix needed to write the null hypothesis associated with the F-test for Alcohol effects in the form $H_0 : \mathbf{C_2}\boldsymbol{\beta} = \mathbf{0}$. (10 marks)

Ans.

$$\mathbf{C_2} = \mathbf{1_a^T} \otimes [\mathbf{I}_{b-1}| - \mathbf{1}_{b-1}] = \mathbf{1_2^T} \otimes [\mathbf{I}_{3-1}| - \mathbf{1}_{3-1}] = \begin{bmatrix}1\\1\end{bmatrix} \otimes \begin{bmatrix}1 & 0| & -1\\ 0 & 1| & -1\end{bmatrix} = \begin{bmatrix}1 & 0 & -1 & 1 & 0 & -1\\ 0 & 1 & -1 & 0 & 1 & -1\end{bmatrix}$$

(e) Present a formula for $SS_{H_{0,2}}$, corresponding to the null hypothesis in part (d), and state it's distribution when the null hypothesis is true.(10 marks)

Ans.

$$\begin{aligned} SS_{H_{0,2}} &= (\mathbf{C_2b-0})^\mathbf{T}[\mathbf{C_2(D^TD)^{-1}C_2^T}]^{-1}(\mathbf{C_2b-0})\\ &= \mathbf{y^TD(D^TD)^{-1}C_2^T[C_2(D^TD)^{-1}C_2^T]^{-1}C_2(D^TD)^{-1}D^Ty} \end{aligned}$$

Under H_0, $SSH_{0,2} \sim \frac{1}{\sigma^2}\chi^2_2$

(f) Compute $SS_{H_{0,2}}$. (10 marks)

Ans.

$$\mathbf{b} = \begin{bmatrix}\bar{Y}_{11.} & \bar{Y}_{12.} & \bar{Y}_{13.} & \bar{Y}_{21.} & \bar{Y}_{22.} & \bar{Y}_{23.}\end{bmatrix}^T = \begin{bmatrix}90.8 & 89.333 & 89.2 & 89.667 & 95.7 & 92.967\end{bmatrix}^T$$

$$\mathbf{C_2b} = \begin{bmatrix}1 & 0 & -1 & 1 & 0 & -1\\ 0 & 1 & -1 & 0 & 1 & -1\end{bmatrix}\begin{bmatrix}90.8\\89.333\\89.2\\89.667\\95.7\\92.967\end{bmatrix} = \begin{bmatrix}-1.7\\2.866\end{bmatrix}$$

$$\mathbf{D^TD^{-1}} = diag(\tfrac{1}{2}, \tfrac{1}{3}, \tfrac{1}{4}, \tfrac{1}{3}, 1, \tfrac{1}{3})$$

$$C_2(D^TD)^{-1}C_2^T] = \begin{bmatrix}1 & 0 & -1 & 1 & 0 & -1\\ 0 & 1 & -1 & 0 & 1 & -1\end{bmatrix}\begin{bmatrix}\frac{1}{2} & 0 & 0 & 0 & 0 & 0\\ 0 & \frac{1}{3} & 0 & 0 & 0 & 0\\ 0 & 0 & \frac{1}{4} & 0 & 0 & 0\\ 0 & 0 & 0 & \frac{1}{3} & 0 & 0\\ 0 & 0 & 0 & 0 & 1 & 0\\ 0 & 0 & 0 & 0 & 0 & \frac{1}{3}\end{bmatrix}\begin{bmatrix}1 & 0\\ 0 & 1\\ -1 & -1\\ 1 & 0\\ 0 & 1\\ -1 & -1\end{bmatrix}$$

$$= \begin{bmatrix} 1.4167 \; 0.5833 \\ 0.5833 \; 1.9167 \end{bmatrix}$$

$$SSH_{0,2} = (\mathbf{C_2 b} - \mathbf{0})^{\mathbf{T}}[\mathbf{C_2}(\mathbf{D^T D})^{-\mathbf{1}}\mathbf{C_2^T}]^{-\mathbf{1}}(\mathbf{C_2 b} - \mathbf{0}) = \begin{bmatrix} -1.7 \; 2.866 \end{bmatrix} \begin{bmatrix} 0.807 & -0.2456 \\ -0.2456 & 0.5965 \end{bmatrix}$$

$$= 9.6252$$